Inharmonicity-Based Method for the Automatic Generation of Guitar Tablature

Isabel Barbancho, *Senior Member, IEEE*, Lorenzo J. Tardón, Simone Sammartino, and Ana M. Barbancho

***Abstract*—In this paper, a system for the extraction of the tablature of guitar musical pieces using only the audio waveform is presented. The analysis of the inharmonicity relations between the fundamentals and the partials of the notes played is the main process that allows to estimate both the notes played and the string/fret combination that was used to produce that sound. A procedure to analyze chords will also be described. This procedure will also make use of the inharmonicity analysis to find the simultaneous string/fret combinations used to play each chord. The proposed method is suitable for any guitar type: classical, acoustic and electric guitars. The system performance has been evaluated on a series of guitar samples from the RWC instruments database and our own recordings.**

***Index Terms*—Fret, guitar, inharmonicity, music analysis, pitch estimation, string, tablature.**

I. Introduction

The guitar is one of the most popular instruments in the world. The guitar is played by professional musicians but it is also well suited for people who like music and want to play music on their own even without any knowledge of music theory. There are different guitar types but most of them have six strings. The length of modern guitar strings is about 65 cm and they are tuned to the following notes: $E2 = 82$ Hz, $A2 = 110$ Hz, $D3 = 147$ Hz, $G3 = 196$ Hz, and $E4 = 330$ Hz [2]. Depending on the guitar type, strings can be made of nylon or steel. In this paper, we will use the term classical guitar for guitars with nylon strings and acoustic guitar for guitars with steel strings that are not electric guitars. Each string of a guitar is slightly different to the others, not only the diameter changes, but also the way in which the strings are made is different [3].

With a guitar, the same pitch can be played using different string, fret, and finger combinations. In this paper, we will deal with the string/fret combinations that can be used to play the different notes with classical, acoustic and electric guitars with six strings. Nevertheless, the method and the ideas developed can be extended to other guitars with any number of strings.

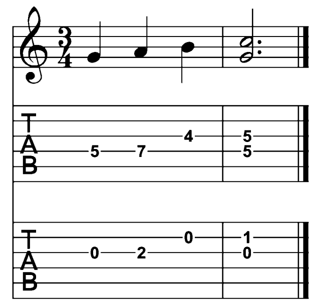

Fig. 1. Score and two possible tablatures to play that score.

There are two main possibilities to write music for guitar: the classical score can be used but also the guitar tablature can be employed. These two choices contain different levels of information on timing, fretting, and fingering [4]. Fig. 1 shows an example of a guitar score and two possible tablatures for that score. The tablature for playing a certain score is not unique. In the tablature, each line represents a string of the guitar. The position of the number indicates the string that must be plucked and the number itself indicates the fret number in the guitar neck where the corresponding string must be pressed with a finger of the left hand to produce the desired note when plucking the string with the right hand. Note that the number increases as the length of the vibrating part of the string is shortened. The numbering in the tablature starts with 0 to indicate that the string must not be pressed at any fret of the guitar neck. As it can be observed, the same notes can be played in different strings of the guitar. Even when the played notes are the same, the sound will be slightly different and also the difficulty of the performance changes. Top guitar players seem to be able to select better string/fret combinations to play than average players and their choices define a determined sonority for the guitar pieces they play and a convenient fingering.

As it can be observed in Fig. 1, the classical score does not contain information about fretting and fingering [5]. Therefore, in order to play a musical piece with a guitar using a score, it is necessary to have some knowledge on music theory to read the score and select, for each note, a correct string/fret combination and to decide the finger of the left hand that must be employed. On the other hand, guitar tablature is a schematic representation of the guitar fingerboard, where the representation of the fretboard positions substitute the notes and the numbers identify the frets. Although tablature does not include timing information, it has become a widely spread music notation format on the Internet [6] maybe because of its readability but also because it can be represented in text format, without graphics.

Manuscript received March 18, 2011; revised October 17, 2011 and January 07, 2012; accepted February 29, 2012. Date of publication March 19, 2012; date of current version April 25, 2012. This work was supported by the Ministerio de Ciencia e Innovación of the Spanish Government under Project TIN2010-21089-C03-02 and Project IPT-2011-0885-430000. The associate editor coordinating the review of this manuscript and approving it for publication was Mr. James Johnston.

The authors are with the Department of Communications Engineering, ETSI Telecommunications, University of Malaga, Malaga E29071, Spain (e-mail: ibp@ic.uma.es; lorenzo@ic.uma.es; ssammartino@ic.uma.es; abp@ic.uma.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASL.2012.2191281

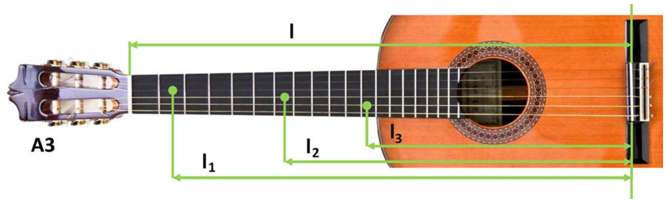

Fig. 2. Three possibilities to play the note $A3$ with a classical guitar.

In this paper, a novel method to determine the tablature used to perform a piece of guitar music using the waveform of the sound of the musical piece is presented. With such a system, we can recover information about the fretting selected by top guitar players in their recordings. Note that a system that is capable of automatically generating the tablature of a guitar piece is able to detect not only the pitch of the played notes, but also the strings and the frets involved. The system we propose in this contribution works with both single notes and chords played with any type of guitar (classical, acoustic and electric guitars). The proposed method is based on the analysis of the behavior of the inharmonicity of guitar strings. The inharmonicity of strings refers to the amount that separates the actual mode frequencies from a harmonic series [7] and it depends on the behavior of the material the string is made of, its radius, its tension and its length. Due to the inharmonicity of the string vibration, the second harmonics of the notes in a certain octave are slightly sharper than the fundamental of the notes in the next octave up and a similar deviation affects the higher harmonics.

Recall that the same note can be played in different strings of a guitar. Fig. 2 shows the three possibilities (fretting positions) to play the note $A3$. The inharmonicity will be different in each of these possibilities because in each of them, the length, the tension, the radius, and maybe the material of the played string will be different. The difficulty stems from the measurement of the inharmonicity of each string and the use of this information to determine both the pitch and the played strings, because the frequency deviations are quite small and the magnitude of the inharmonicity coefficients is very small.

The method proposed in this paper achieves very good performance in the process of determination of both the pitch and the played strings by means of inharmonicity analysis. As it will be shown, the best performance is obtained if the inharmonicity coefficients are estimated before running the tablature generation system, although predefined sets of coefficients can also be used. The method is used to analyze not only single notes but also chords. As it will be clear from the literature review, we have found no references in which the tablature is obtained using the audio waveform of a guitar piece. In a previous work [8], we presented a system to detect the pitch and played string for single guitar notes. In that case, the proposed system used a set of time and frequency features extracted from the guitar sound (frequency centroid, nontonal spectral pattern [9], exponential decay of harmonics, etc.), but it did not attain good results in the determination of the played string.

In the literature, there is a considerable amount of references that deal with the problem of detection of the notes played in a certain musical piece, i.e., transcription systems. In the paper by Klapuri [10], the concepts of harmonicity and spectral smoothness are used for the detection of multiple fundamental frequencies on the recordings of different instruments. Yeh [11] presents a method for the estimation of multiple fundamental frequencies of polyphonic music signals based on the short-time Fourier transform. Other authors deal with the transcription of specific instruments: drums, piano, etc. Focusing on the guitar, in the paper by Gagnon *et al.* [12], the authors propose a system based on neural-networks to detect the number of strings that sound in a chord played with a guitar and, using this information, they propose some hand positions to play that sound but they do not solve the ambiguity of the problem. In the paper by O'Grady and Richard [13], a system for hexaphonic guitar transcription is presented but, in this case, the guitar is modified so that they are able to separately record the sound of each of the six guitar strings.

Looking for references on inharmonicity analysis in string instruments and its application, we found that Järveläinen *et al.* study mainly the audibility of the inharmonicity and its importance for digital sound synthesis [14], but they do not use it to determine the played string in a guitar. Godsill and Davy [15] propose a method for modeling the inharmonicity in musical instruments, although it is not focused for any particular application. Other studies have also tackled inharmonicity estimation. For example, Galembo and Askenfelt [16] present different techniques for estimating the inharmonicity of piano strings and Rauhala *et al.* [17] present an iterative algorithm for the estimation of the inharmonicity coefficient of slightly inharmonic stringed instrument sounds.

There are also references dealing with the tablature in guitar music, but they do not use the audio waveform of the guitar. In order to obtain the tablature, they use either the guitar score [4], [18]–[21] or image-processing systems [22]–[24], but the recorded waveform of the guitar sound is not used, unlike the procedure that we propose in this contribution. Also, some references can be found in which optical tablature recognition (OTR) is used to generate a MIDI output [25]. Only Traube and Smith [26] made an attempt to obtain the fingering point using the audio waveform of a guitar piece. They describe a method to obtain the plucking point using the general equation of a vibrating string of a certain length with fixed ends. There, they explain that under the hypothesis that the right-hand fingers pluck the strings in a narrow area close to the tone hole, the plucking point can be used to determine the fingering. Unfortunately, they do not show any result of the performance of the fingering estimation procedure, although they explain that the performance of the method would depend on the accuracy of the estimation of the plucking point. In other works [27]–[29], only the plucking point is estimated using a weighted least-square estimation of a comb filter delay.

In this paper, a complete work flow for the extraction of the tablature of single guitar notes or chords using the waveform of the guitar sound is detailed. The outline of the paper is as follows: after a brief description of the main features of the guitar, in Section II, the main theoretical aspects of the work are described in Sections III–V. The method for the estimation of the note/string/fret combination is detailed in Section VI. In

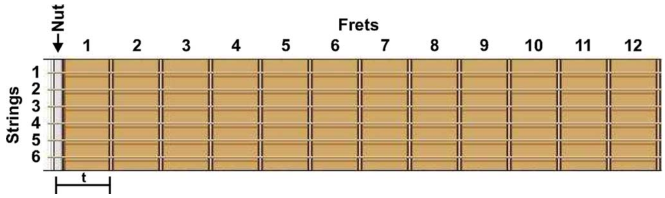

Fig. 3. Schematic representation of the guitar fingerboard.

Section VII, a procedure is developed for the analysis of chords. Results on the evaluation of the performance of the method proposed are presented in Section VIII and, finally, some conclusions are drawn in Section IX.

II. Main Features of the Guitar

The guitar is a plucked string instrument basically composed of a resonance body, a fingerboard hosting the frets, and a headstock where the tuning keys of the strings are fixed. The tension of the strings is tunable by the keys. In Fig. 3, a schematic representation of the guitar fingerboard is presented in which strings and frets are labeled with their corresponding numbers.

The guitar needs the concurrent use of the two hands to play a note (except when the open strings are played). Usually, the left hand presses one or several strings on the corresponding frets and the right hand plucks the selected strings. For each string at each fret pressed a single note sounds and between any fret and the next one there is a separation of a semitone. Note that the same note can be played in up to three different string/fret pairs.

It is well known that the ratios between the main frequencies $f_1(s)$ and $f_2(s)$ of two sounds played with the same string s and the corresponding lengths of the vibrating part of the strings l_1 and l_2, respectively, are inversely proportional:

$$\frac{f_1(s)}{f_2(s)} = \frac{l_2}{l_1} \tag{1}$$

Now, let t represent the distance between the guitar nut and the first fret (Fig. 3). Then, the ratio between the fundamental frequency of the sound played when a string s is pressed at the first fret $f_0(s,1)$ and the main frequency found with the open string $f_0(s,0)$ can be written as follows:

$$\frac{f_0(s,1)}{f_0(s,0)} = \frac{l}{l-t} = 2^{\frac{1}{12}} \tag{2}$$

where l is the length of the strings as shown in Fig. 2. Therefore, $(l-t)$ is the portion of the string vibrating when $f_0(s,1)$ sounds. The term $2^{1/12}$ is related to the ratio of the frequencies of two subsequent frets on the guitar fingerboard (two subsequent notes in the twelve-tone equal temperament scale) and it is due to the way in which guitars are built, regardless the tuning of each of the strings [30]. Hence, in general:

$$f_0(s,n) = f_0(s,0) \cdot 2^{\frac{n}{12}} \tag{3}$$

where n is the fret number.

Then, (2) can be rewritten as follows:

$$t = l \cdot \left(1 - \frac{1}{2^{\frac{1}{12}}}\right). \tag{4}$$

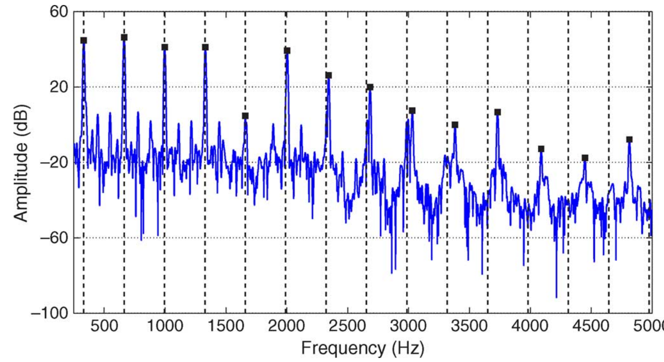

Fig. 4. Spectrum of the note E4, with $f_0(3,9) = 329.63$ Hz. Detected partials are marked with filled squares. Multiples of the fundamental frequency are marked using dashed vertical lines.

These relations will be used later to define inharmonicity relations.

III. Inharmonicity of the Partials

The inharmonicity of the partials of string instruments is an important topic that is being discussed on the basis of the human perception [31]. From an analytic point of view, the partials are a crucial feature of the spectrum of the waveform of the sound of musical instruments.

Each note played plucking the string s pressed at fret n can be seen as the combination of its corresponding fundamental frequency component, $f_0(s,n)$, and a series of partials whose frequencies $f_k(s,n)$ are, ideally, multiple of $f_0(s,n)$:

$$f_k(s,n) = k \cdot f_0(s,n). \tag{5}$$

A shift of the partials of the spectrum is the outcome of the inharmonicity of the instrument [32]. Fig. 4 displays the spectrum of the note E4, played on the 9 th fret of the 3rd string, with fundamental frequency $f_0(3,9) = 329.63$ Hz. Observe the relative position between the partials detected $\hat{f}_k(3,9)$ (marked with black filled squares) and the theoretical expected frequencies $f_k(3,9)$ (dashed vertical lines) shown in the figure.

The expected frequency of the partials taking into account the inharmonicity, are separated from the multiples of $f_0(s,n)$ a quantity that depends on the so called inharmonicity coefficient (β), as described in the following equation:

$$f_k^*(s,n) = k \cdot f_0(s,n) \cdot \sqrt{1+\beta \cdot k^2} \tag{6}$$

where the expected frequencies of the partials, $f_k^*(s,n)$, are related to the corresponding kth multiples of the fundamental frequency $f_0(s,n)$, but a term that depends on the inharmonicity coefficient of the played string β is involved in the relation. This coefficient β depends on the behavior of the material used for the string, on the string diameter d, its length l and its tension T:

$$\beta = \frac{\pi \cdot Q \cdot d^4}{64 \cdot l^2 \cdot T} \tag{7}$$

where the term Q, called Young's modulus, is a constant related to the resistance of elastic bodies to deformation [32]. It is important to observe that the measured frequencies of the partials are always higher than the corresponding multiples of $f_0(s,n)$ [32]. Also, the separation between them increases with the number of the partial involved, k.

IV. Estimation of the Inharmonicity Coefficient

In order to be able to estimate the note played and the string and the fret involved, the inharmonicity coefficient must be accurately estimated. Taking into account (7), it is easy to come to the conclusion that there is one β coefficient for each combination of string s and fret n, $\beta(s,n)$.

Using the parameter t defined in (4) and (7), we can write the following relation:

$$\begin{aligned}\beta(s,1) &= \frac{\pi \cdot Q \cdot d^4}{64 \cdot (l-t)^2 \cdot T} = \frac{\pi \cdot Q \cdot d^4}{64 \cdot l^2 \cdot T \cdot 2^{-\frac{1}{6}}} \\ &= \beta(s,0) \cdot 2^{\frac{1}{6}} \end{aligned} \tag{8}$$

where the tension T is assumed to be constant.

At the sight of (8), and focusing on guitar strings, it is found that the coefficient $\beta(s,n)$ for a specific string s and fret n is the product of the inharmonicity coefficient of the preceding fret, $\beta(s,n-1)$, multiplied by $2^{1/6}$. This consideration allows to derive the inharmonicity coefficient of a string at any fret with only one known $\beta(s,n)$ of that string:

$$\beta(s,n) = \beta(s,0) \cdot 2^{\frac{n}{6}} \tag{9}$$

where $\beta(s,0)$ is the inharmonicity coefficient of the open string s.

A. Deviation of the Partials

The separation between the ideal frequency of a partial (a multiple of $f_0(s,n)$) neglecting inharmonicity and its expected measured value using our inharmonicity model $f_k^*(s,n)$, is defined as frequency deviation. The expected frequency deviation of the kth partial can be obtained subtracting (5) from (6):

$$\begin{aligned} d_k(s,n) &= f_k^*(s,n) - f_k(s,n) \\ &= k \cdot f_0(s,n) \cdot \sqrt{(1+\beta(s,n) \cdot k^2} - k \cdot f_0(s,n). \end{aligned} \tag{10}$$

Replacing the terms $f_0(s,n)$ and $\beta(s,n)$ in (10) by the relations in (3) and (9), respectively, we find an expression of the frequency deviation parameterized by the number of the partial (k) and the fret number n for any string s:

$$d_k(s,n) = k \cdot f_0(s,0) \cdot 2^{\frac{n}{12}} \cdot \sqrt{\left(1+\beta(s,0) \cdot 2^{\frac{n}{6}} \cdot k^2\right)} - 1. \tag{11}$$

Equation (11) clearly reveals that the frequency deviation increases with the number of partial k and with the inharmonicity coefficient $\beta(s,0)$. Also note that if $\beta(s,0) \rightarrow 0$, then the partials tend to be located at the corresponding multiple of $f_0(s,n)$.

B. Dealing With the Matching Error

The Fourier transform is used to extract the spectral information of the signal under analysis. However, a matching error is found in the analysis of the inharmonicity coefficient due to the error in the estimation of any frequency value due to the effect of noise, windowing and the limited frequency resolution of the discrete Fourier transform and the goodness of the fit between the frequencies of the fundamental and partials and the inharmonicity model. This matching error can negatively affect the procedure of estimation of the inharmonicity coefficient, so we develop a model to take into account this issue to obtain more reliable estimates.

Consider the hypothesis according to which the measured fundamental frequency of a note $\hat{f}_0(s,n)$ differs from its real value $f_0(s,n)$ a quantity $\epsilon_0(s,n)$ such that

$$\hat{f}_0(s,n) = f_0(s,n) + \epsilon_0(s,n). \tag{12}$$

This means that the expected frequencies of the partials $f_k^*(s,n)$ (6) will be affected by this error:

$$f_k^*(s,n) = k \cdot (f_0(s,n) + \epsilon_0(s,n)) \cdot \sqrt{1+\beta(s,n) \cdot k^2}. \tag{13}$$

Observe that the influence of the error in $f_k^*(s,n)$ increases with the number of partial.

Not only the fundamental frequency is affected by a measurement error but also the measured partials $\hat{f}_k(s,n)$ are affected. As a consequence, we can consider a similar model for the measured partials:

$$\hat{f}_k(s,n) = k \cdot f_0(s,n) \cdot \sqrt{1+\beta(s,n) \cdot k^2} + \epsilon_k(s,n). \tag{14}$$

At this stage, we build a model for the frequency deviation of the partials due to the inharmonicity [see (10)], taking into account the following measurement errors:

$$\begin{aligned} c_k(s,n) &= \hat{f}_k(s,n) - k \cdot \hat{f}_0(s,n) \\ &= k \cdot f_0(s,n) \cdot \left(\sqrt{1+\beta(s,n) \cdot k^2} - 1\right) \\ &\quad - k \cdot \epsilon_0(s,n) + \epsilon_k(s,n). \end{aligned} \tag{15}$$

Note that (15) takes into account the measurement error of both the fundamental frequencies and the partials. Using (15) we develop a polynomial model to fit $c_k(s,n)$. Making use of the fact that the magnitude of the parameter β to be estimated is very small, it is possible to use an approximation of $\sqrt{1+\beta \cdot k^2}$ with the first two terms of the Taylor series [33]. Then, (15) can be written as follows:

$$\begin{aligned} c_k(s,n) &\cong k \cdot f_0(s,n) \cdot \left(1 + \frac{1}{2} \cdot \beta(s,n) \cdot k^2 - 1\right) \\ &\quad - k \cdot \epsilon_0(s,n) + \epsilon_k(s,n) \\ &= \frac{1}{2} \cdot \beta(s,n) \cdot f_0(s,n) \cdot k^3 \\ &\quad - k \cdot \epsilon_0(s,n) + \epsilon_k(s,n). \end{aligned} \tag{16}$$

Now, a least squares method can be employed to fit a third degree polynomial function $y = a \cdot k^3 + b \cdot k + c$ to the data in order to find the coefficients a and b in this model. Finally, the inharmonicity coefficient $\beta(s,n)$ can be estimated using

$$\beta(s,n) = \frac{2 \cdot a}{\hat{f}_0(s,n) + b}. \tag{17}$$

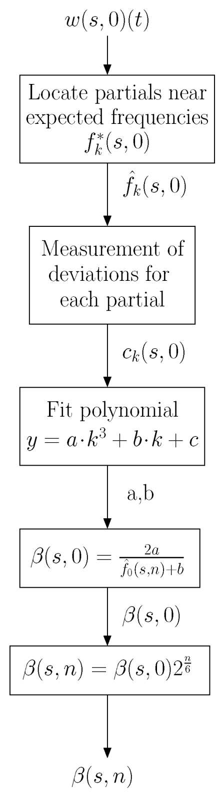

Fig. 5. Diagram for the estimation of the inharmonicity coefficient.

C. *Estimating the Inharmonicity Coefficient in Practice*

The inharmonicity coefficient must be estimated from the signal of any note played in the guitar, assuming that its fundamental frequency is known. In our system, the inharmonicity coefficient of the open string, $\beta(s,0)$ will be calculated. Recall that taking into account (9), the inharmonicity coefficients $\beta(s,n)$ for each fret n are easily determined.

The algorithm to obtain $\beta(s,n)$ is depicted in Fig. 5. First of all, the Fourier transform of the audio waveform $w(s,0)(t)$ corresponding to the note played with the open string s is calculated. Then, the partials are located using a series of windows centered at $k \cdot f_0(s,0)$. The frequency deviation of each partial is calculated and the function $y = a \cdot k^3 + b \cdot k + c$, with k the number of partial, is fitted to estimate the parameters a and b. Afterward, $\beta(s,0)$ is estimated using (17) and, finally, $\beta(s,n)$ are defined using (9).

In order to improve the accuracy of the estimation increasing the number of partials detected, the algorithm is executed iteratively taking into account the inharmonicity coefficient obtained at each new iteration. At each stage, the windows are centered at the frequencies suggested by the function approximated at the previous step. Up to 50 partials are searched although this number is also limited according to the sampling rate.

An illustration of the evolution of the algorithm used to obtain the inharmonicity coefficient of the note E4 is shown in Fig. 6. The first set of measured partials obtained at the first iteration is used to define the first estimation of the inharmonicity coefficient. Then, at each iteration, the windows are shifted according to the inharmonicity coefficient estimated at the previous step. The width of the windows is always half the fundamental frequency.

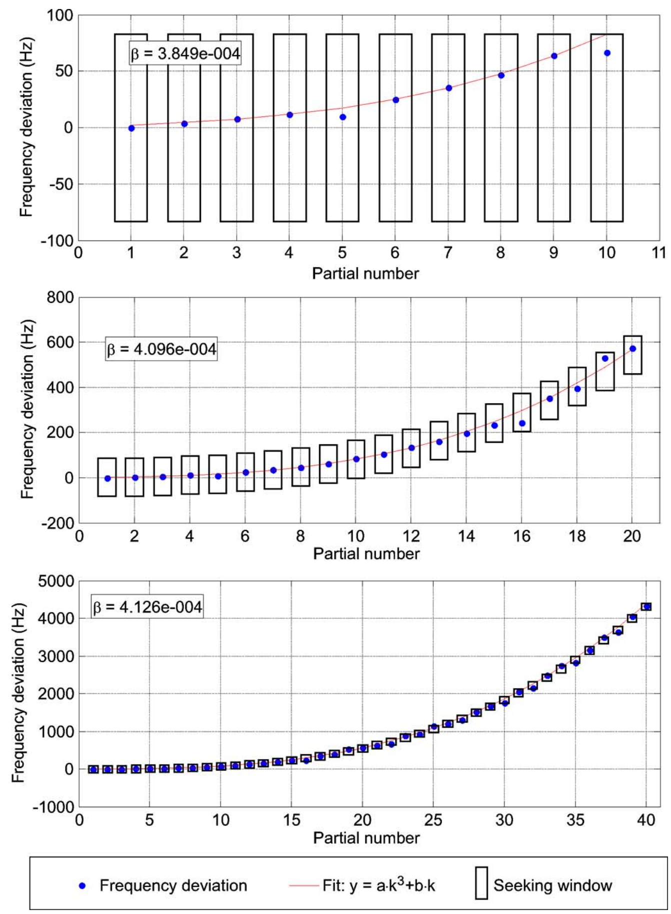

Fig. 6. Illustration of the refinement of the calculus of the inharmonicity coefficient for note E4, played on the 9th fret of the 3rd string. Iterations 1, 2, and 3 are shown from top to bottom, respectively. The search windows are shown as boxes. The measured deviations for each partial are marked with dots. The fitted polynomial at each iteration is also drawn.

The iterative procedure ends when the estimated $\beta(s,n)$ at a certain iteration changes less than 1% with respect to the value found at the previous iteration.

V. Matching Error of the Fundamental Frequency and Partials

The computation of the matching error between our inharmonicity model and all the possible combinations of the measured fundamental frequencies and partials for each possible note will be used in the process of detection of the string/fret combination used to play each note. Our model makes use of the matching error of both the fundamental frequency, $\epsilon_0(s,n)$, and partials, $\epsilon_k(s,n)$. Their relation can be estimated comparing the measured partials (directly affected by $\epsilon_k(s,n)$) and the expected locations of the partials calculated using the measured fundamental frequency (these frequencies are indirectly affected by $\epsilon_0(s,n)$).

To begin with, observe that using (13) and (14), we can compare the measured frequencies of the partials $\hat{f}_k(s,n)$ with their expected values $f_k^*(s,n)$, under the hypothesis that the measured critical frequencies (fundamental and partials) correspond to a played note:

$$f_k^*(s,n) - \hat{f}_k(s,n) = \epsilon_0(s,n)\cdot k\cdot\sqrt{1+\beta(s,n)\cdot k^2} - \epsilon_k(s,n). \tag{18}$$

Recall that it is the note/string/fret combination what we want to determine. To this end, we must take into account different inharmonicity coefficients, in particular, the coefficient $\beta^*(s,n)$ related to the expected partials $f_k^*(s,n)$. Note that it must be checked whether these partials correspond to a fundamental frequency or not. Also, the inharmonicity coefficient $\hat{\beta}(s,n)$ must be taken into account. This parameter is related to the measurement of the partials $\hat{f}_k(s,n)$. In order to consider all the possibilities, we assume that, in the general case, these two inharmonicity parameters can be different. In this way, we take into account that different strings can be used to play a note. Then, (18) must be rewritten as follows:

$$\begin{aligned} f_k^*(s,n) &- \hat{f}_k(s,n) \\ &= k\cdot f_0(s,n)\cdot\left(\sqrt{1+\beta^*(s,n)\cdot k^2} - \sqrt{1+\hat{\beta}(s,n)\cdot k^2}\right) \\ &\quad + \epsilon_0(s,n)\cdot k\cdot\sqrt{1+\beta^*\cdot k^2} - \epsilon_k(s,n). \end{aligned} \tag{19}$$

Equation (19) can be divided by $[k\cdot\sqrt{1+\beta^*(s,n)\cdot k^2}]$ to obtain

$$\begin{aligned} \frac{f_k^*(s,n)-\hat{f}_k(s,n)}{k\cdot\sqrt{1+\beta^*(s,n)\cdot k^2}} &= f_0(s,n)\cdot\left(1-\frac{\sqrt{1+\hat{\beta}(s,n)\cdot k^2}}{\sqrt{1+\beta^*(s,n)\cdot k^2}}\right) \\ &\quad + \epsilon_0(s,n) - \frac{\epsilon_k(s,n)}{k\cdot\sqrt{1+\beta^*(s,n)\cdot k^2}}. \end{aligned} \tag{20}$$

The terms $\sqrt{1+\beta(s,n)\cdot k^2}$, with $\beta = \{\beta^*, \hat{\beta}\}$, in (20) can be approximated by the first two terms of the Taylor series for $\sqrt{1+p}$, as in (15). Hence, the term $[1 - (\sqrt{1+\hat{\beta}(s,n)\cdot k^2}/\sqrt{1+\beta^*(s,n)\cdot k^2})]$ in (20) can be rewritten as follows:

$$1-\frac{\sqrt{1+\hat{\beta}(s,n)\cdot k^2}}{\sqrt{1+\beta^*(s,n)\cdot k^2}} \cong \frac{1}{2}\cdot\left(\beta^*(s,n)-\hat{\beta}(s,n)\right)\cdot k^2 \tag{21}$$

where we have also used the following: $\sqrt{1+\beta^*(s,n)k^2} \approx 1$.

Using this expression in (20), we obtain

$$\begin{aligned} &\frac{f_k^*(s,n)-\hat{f}_k(s,n)}{k\cdot\sqrt{1+\beta^*(s,n)\cdot k^2}} \\ &\cong f_0(s,n)\cdot\frac{1}{2}\cdot\left(\beta^*(s,n)-\hat{\beta}(s,n)\right)\cdot k^2 + \epsilon_0(s,n) - \frac{\epsilon_k(s,n)}{k}. \end{aligned} \tag{22}$$

Now, we assume that the influence of $\epsilon_k(s,n)$ becomes negligible for partials higher than 9. Then, the following approximation is found:

$$\begin{aligned} \frac{f_k^*(s,n)-\hat{f}_k(s,n)}{k\cdot\sqrt{1+\beta^*(s,n)\cdot k^2}} &\cong f_0(s,n)\cdot\frac{1}{2}\cdot\left(\beta^*(s,n)-\hat{\beta}(s,n)\right)\cdot k^2 \\ &\quad + \epsilon_0(s,n), \text{ with } k > 10. \end{aligned} \tag{23}$$

Observe that this is a quadratic function in which the coefficient of the quadratic term depends on the difference between the two inharmonicity coefficients $\beta^*(s,n)$ and $\hat{\beta}(s,n)$. In particular, this function becomes a constant, $\epsilon_0(s,n)$, when the two coefficients are the same, that is to say, the partials measured and the partials calculated on the basis of the measured fundamental frequency are affected by the same inharmonicity coefficient. When this condition is fulfilled, it means that we have found a coherent set of measured critical frequencies (fundamental and partials) that correspond to a note played plucking a certain string. We say that (23) represents an approximation of the matching error of the fundamental frequency related to the measured partials. This expression is of main importance in the detection process that will be explained in the next section.

VI. Detection of Note/String/Fret Combinations

In this section, the formulas derived previously will be used to determine the note played, the string plucked and the fret where the string was pressed to play that note using solely the audio waveform.

Since the played note is unknown, the fundamental frequency must be found. To this end, a set of candidate fundamental frequencies of the note played is found using a normalized version of the magnitude of the spectrum of the signal. The spectrum is normalized so that its maximum amplitude is 1. Then, all the components in this normalized spectrum with amplitude under 0.01 (1% of the highest peak) are zeroed and the frequency range of the fundamental frequencies (from 80 Hz to 660 Hz) of all the notes that can be played with a guitar (from the note E2 on the open 6th string to the note E5 on the 12th fret on the 1st string) is analyzed. In this range, all the local maxima are located and the ones that are closer to any of the possible fundamental frequencies are selected to define the candidate notes.

Then, all the possible string/fret combinations for all the candidate notes are considered. Also, their corresponding inharmonicity coefficients (see Section IV) are employed in the following analysis to make use of all the sets of candidate partials that can be defined.

Recall that for each candidate note (fundamental frequency) a series of up to three possible string/fret combinations can be defined, i.e., the number of strings involved in the analysis of each note can be three at most (Section II). For each string/fret combination, we calculate the inharmonicity coefficient $\beta(s,n)$ as explained in Section IV-C. Then, a series of windows (frequency bands) is defined on the basis of the inharmonicity coefficient. Up to 50 windows are defined. Partials are sought in these windows. The largest peak in each of these windows is considered to be a candidate partial. The windows are centered at the values of the calculated partials $f_k^*(s,n)$ (13) and their width grows with the number of partial k. Thus, each window $v_k(s,n)$, is defined by the following frequency range:

$$v_k(s,n) \rightarrow k\cdot\sqrt{1+\beta(s,n)\cdot k^2}\cdot\left(\hat{f}_0(s,n) \pm R\right) \tag{24}$$

where R should be large enough to capture all the possible partials ($R > |\epsilon_f|$), but not too large to avoid and excess of outliers and spikes. A value of $R = 10$ Hz has been employed at this stage (see Fig. 7).

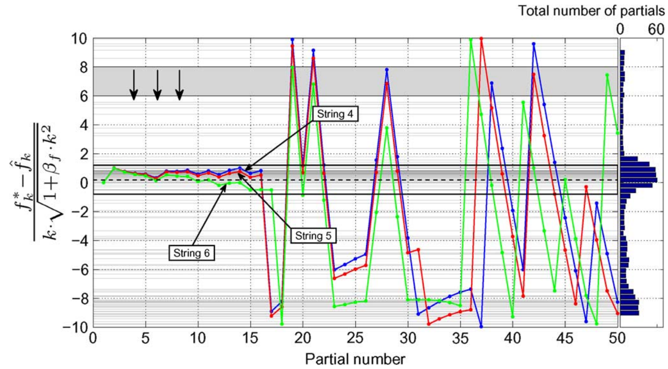

Fig. 7. Example of the voting procedure for the estimation of the error ϵ_f for the note E3 using $v_k(s,n)$ as defined in (24), with $R = 10$ Hz. Sample containers are shown in gray with dashed edges. The most populated container is displayed in gray with its center marker with a dashed thick line.

Using the partials $\hat{f}_k(s,n)$ found in $v_k(s,n)$, according to (24), for each possible string/fret combination and for each partial found, we calculate the left-hand side of (23). In Fig. 7, this result is plotted versus the partial number k for a candidate note E3.

Now, we use a voting scheme to find the first estimation $\epsilon_0^*(s,n)$ of the error $\epsilon_0(s,n)$. Consider the set of windows defined previously, $v_k(s,n)$, covered by a sliding container of 2 Hz of width (see the gray area at the top of Fig. 7). Now, we sum the number of hits at each position of the container, i.e., we sum the number of partials for all the possible string/fret combinations for each fundamental such that $\epsilon_k(s,n)$ lays in the container. Then, the center of the container at the location with the largest count is taken as the first estimation $\epsilon_0^*(s,n)$ of $\epsilon_0(s,n)$. In Fig. 7, the container that gives the first estimation of $\epsilon_0(s,n)$ is centered at 0 (the right-hand side of this figure represents the sums of hits in the container versus its position). Note that the largest count of the number of partials in the sliding container for candidate notes that were not actually played is usually quite smaller than the largest count for any of the possible string/fret combinations for the note played.

This first estimation needs to be improved. Specifically, we must reduce the width of the windows to search for the partials in order to lower the number of outliers and spikes taken as partials. To this end, we use a voting procedure like the one described, after some adjustments. To begin with, the windows $v_k(s,n)$ are shifted D Hz according to the $\epsilon_f^*(s,n)$ estimated:

$$v_k'(s,n) \to k \cdot \sqrt{1+\beta(s,n)\cdot k^2} \cdot \left(\hat{f}_0(s,n) + D \pm R\right) \quad (25)$$

where $D = -\epsilon_0^*(s,n)$. As before, the following condition should be fulfilled: $R > |\epsilon_0(s,n) + D|$, in order not to loose partials that may fall out of the search windows. At this stage, R is defined to be $R = 2.3$ Hz (see Fig. 8).

Now, we search for candidate partials in $v_k'(s,n)$ for each possible string/fret combination for the notes under analysis. Recall that only the right combination leads to the approximation of the left-hand side of (23) to a constant. At this stage, a container of 0.5 Hz of width is used. Again, the location of the container with the largest number of hits of $\epsilon_f(s,n)$ is selected. Fig. 8 illustrates this process for the example shown in Fig. 7.

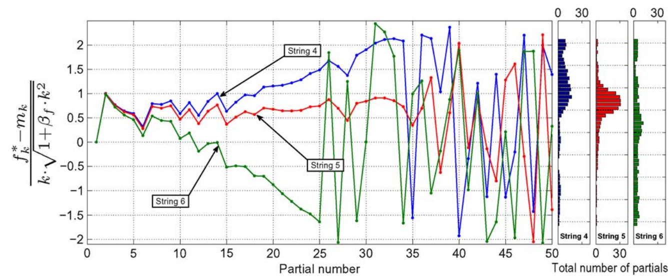

Fig. 8. Example of the second stage of the voting procedure for the detection of the note/string/fret corresponding to a note E3 played, using $v_k'(s,n)$ as defined in (25), with $R = 2.3$ Hz.

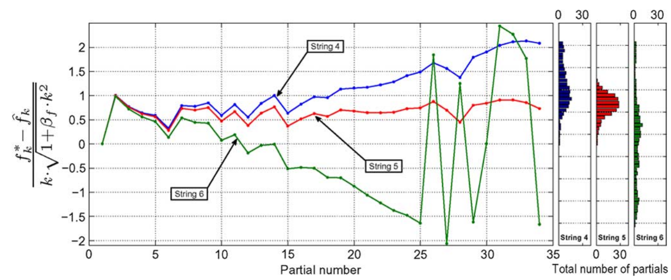

Fig. 9. Example of reduction of the number of partials to search for in the running case (E3 played note). The new maximum number of partials is 34.

The string/fret combination with the largest number of partials in the selected container can be used to identify the note played and the string/fret combination used to play that note. In our example, an E3 note was played using the string (number 5) detected (Fig. 8). 31 partials were found for the winning combination (Fig. 10).

However, we have found that it is possible to enlarge the differences between the candidate note/string/fret combinations and the correct combination in our voting scheme and, at the same time, remove some sparse errors. These improvements can be obtained by simply adapting the maximum number of partials to find. Specifically, consider the winning combination found at the previous stage. Then calculate the 80% of the cardinal of the number of partials of the count of the winning combination, P_c (25 in our example). Afterward, find the highest partial in this selection so that the count of the number of partials in the sliding windows at the location found at the previous stage is P_c and track the partials (see string 5 in Fig. 8) until a measure lays out of the sliding window. The last (highest) partial in the sliding window is used to set the upper limit to the number of partials that could be found in the next stage K_c. In our example, this limit is 34 (see Figs. 8 and 9).

Now, the voting procedure is executed again with $R = 10$ Hz and, then, with $R = 2.3$Hz, but in this case, $k = \{1, \ldots, K_c\}$. In Fig. 10, a schematic representation of the whole estimation process developed to detect the note played and the string/fret combination used to play that note is shown. The data correspond to the running example in which the note E3 was played.

Fig. 10 illustrates how the differences between the correct note/string/fret combination and the other candidate combinations evaluated are enlarged after the imposition of an adaptive limit to the number of partials to analyze. As shown in Fig. 10, the note/string/fret combination that achieves the largest count of the number of partials in our scheme corresponds to the 7th

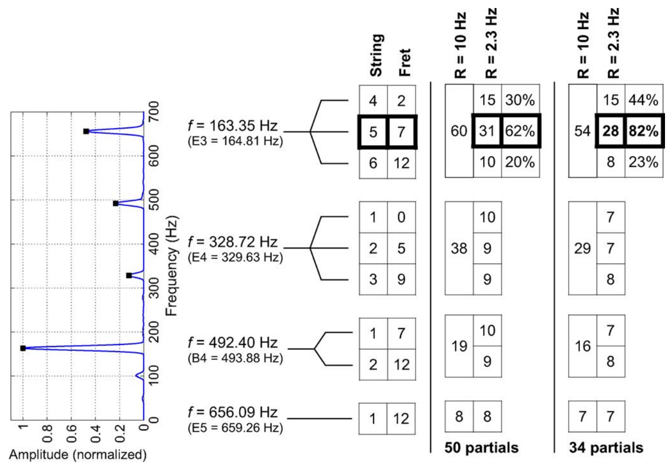

Fig. 10. Schematic representation of the whole procedure to estimate the note played and the string and fret involved. The correct combination of note/string/fret is highlighted in bold.

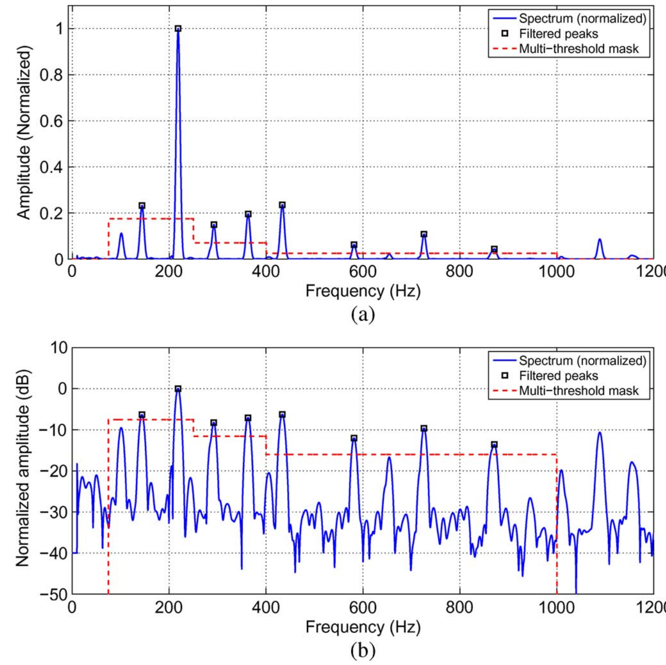

Fig. 11. Example of the application of a threshold to the magnitude of the spectrum of the audio waveform of a guitar chord. (a) The masked spectrum (normalized amplitude). (b) The masked spectrum (normalized amplitude dB).

fret pressed on the 5th string to play an E3 note. This combination gave a count of 31 partials at the end of the first round (62% of hits in the sliding window, with a maximum of 50 partials) and 28 partials (82% of hits with a maximum of 34 partials) at the end of the second round.

VII. Chord Detection

The system described can be applied to estimate groups of notes played at the same time (chords). To this end, two different approaches are considered, depending on the number of notes in the chord.

If the chord can be freely formed by any combination of notes, then, the maximum number of notes to detect is limited to four. On the other hand, fixed combinations or templates can be considered for all the possible chords (barre chords) then, the maximum number of notes that can be detected is six. The scenario to apply the procedure (up to four notes per chord or predefined chord types) must be selected beforehand.

The approaches are based on the estimation of the notes and on the iterative deletion (removal) of the fundamental frequency and partials from the spectrum.

In this scenario, two main problems have to be faced: the possibility of the presence of notes that (approximately) share the partials (equal notes played at different octaves) and the interference or overlap that may happen between partials and fundamental frequencies of different notes.

We will consider first the problem of detection of free chords with four notes or less and, later, the problem of detection of predefined chords.

A. Free Chords—Up to Four Notes

The detection process described previously can be used to evaluate all the possible combinations of fundamental frequencies and partials that can appear in the spectrum of the waveform of a chord. However, this can be an enormous task from the computational point of view. This fact drives us to develop a procedure to reduce the computational burden.

To begin with, a threshold defined on the basis of the relative magnitude of the peaks detected in the lower part of the spectrum is used. The magnitude of the threshold is defined as 17% of the largest peak found in the range $75\ \text{Hz} < f < 250\ \text{Hz}$, 7% of the largest peak in the range $250\ \text{Hz} < f < 400\ \text{Hz}$ and 2.5% of the largest peak in the range $400\ \text{Hz} < f < 1000\ \text{Hz}$. In this range ($75\ \text{Hz} < f < 1000\ \text{Hz}$) all the peaks, and frequency bins, with magnitude under the threshold, are zeroed.

Fig. 11 shows an example of the application of this process. Only the peaks tagged with a square symbol are considered to possibly correspond to a fundamental frequency or partial in the next step. We have not found this process to perform undesired deletions of fundamentals or partials.

Now, it must be checked if the remaining peaks correspond to fundamental frequencies or partials. To this end, we assume that the partials related to each of the fundamental frequencies are separated a distance that is approximately equal to the fundamental frequency. So, the separation in frequency between all the peaks is calculated. If the separation between two peaks is approximately a fundamental frequency, then the highest peak will be considered to be a partial instead of a fundamental frequency and it will be used accordingly, in the next estimation step.

At the following stage, we consider all the possible fundamental frequencies. All the possible partials of all the possible fundamental frequencies are sought and the one with the largest number of partials is selected to be the first note detected (see previous section). Afterwards, the magnitude of the bins that correspond to these frequencies are zeroed to repeat this process iteratively until all the possible fundamental frequencies are considered.

This first estimation (note/string/fret) is considered reliable only if the difference in the number of partials of this estimation and the number of partials of any of the other estimations is larger than 80%. Otherwise, the first estimation is not be considered reliable. In this case, the two most probable string/fret

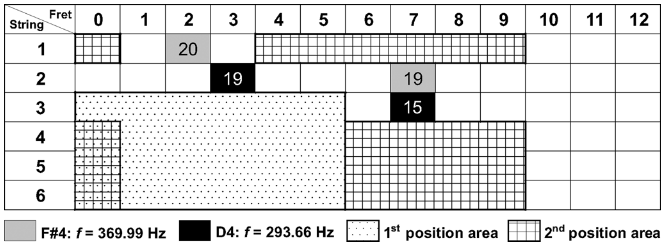

Fig. 12. Example of area delimitation for chord detection. Two possible notes, $F\#4$ and $D4$, are detected at the first stage. All the possible positions of the first two notes are defined with respect of the amount of partials sought. A limiting area is defined around the two possible combinations of positions. Two limiting areas are defined, for the notes played on frets 2 and 3 (dots filled area) and for the notes played on fret 7 (squares filled area). Note that the open string positions are included in both areas showing a partial overlap for the strings 4, 5, and 6.

combinations that can be employed to play the first two notes in the previous count are compared and the combination with the largest sum of partials is considered as the most reliable.

Once the first estimation is considered reliable, then the search for the remaining pitch/string/fret estimations is done according to the following rules: 1) the string used to play the first note is removed from the bunch of possible strings for the remaining estimations (two different notes cannot be played on the same string simultaneously) and 2) the frets of the following estimations are restricted to be in an area around the fret estimated for the first note. This area is centered in the most reliable position.

In Fig. 12, an example of the process and the idea of area delimitation are shown.

The first note $F\#4$ (light gray) has been detected as played on string 1/fret 2 (20 partials found) or on string 2/fret 7 (19 partials found). The second note $D4$ (in black) has been detected as played on string 2/fret 3 (19 partials found) or on string 3/fret 7 (15 partials found). The first note is not readily considered reliable because the difference between the number of partials found for the first string/fret combination and the number of partials found for the first string/fret combination of the second note is smaller than 80%.

Hence, the two first notes positions have to be taken into account. Our procedure determines that the most reliable detection is $F\#4$, since it accounts for a total amount of 39 partials (in this example, 34 partials are found for the note $D4$).

The search area is located around the first position (filled with a dotted pattern). In the example, the third note will be correctly detected as A4 ($f = 220$ Hz), to form the first position of the D major chord.

This procedure requires a set of conditions to stop. The first condition refers to maximum number of simultaneous notes allowed, which is four. The second condition is an adaptive threshold on the number of partials detected for a possible note. If this quantity is under 50% of the number of partials found for the first (the best) detection, then the algorithm ends. Note that a rather similar behavior is expected in the detection of the different notes of a chord, so we have set this conservative limit to the minimum number of partials to consider played notes so as not to loosen actually played notes.

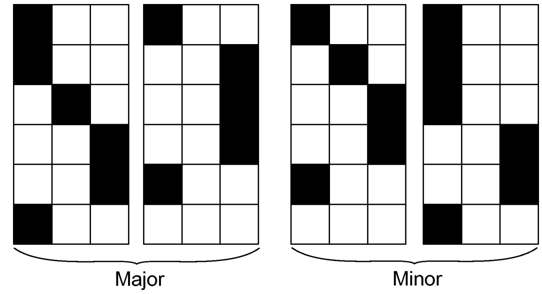

Fig. 13. Four chord templates defined for the two major and minor positions.

TABLE I
EXAMPLE OF DETECTION OF THE CHORD G MAJOR. POSSIBLE FUNDAMENTAL FREQUENCIES

Index	1	2	3	4	5	6	7	8
Pitch(Hz)	96.9	245.9	146.5	194.1	292.9	491.6	585.8	391.1

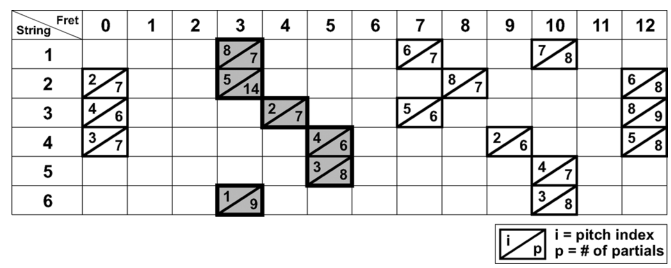

Fig. 14. Illustration of template matching. Evaluation of all the possible template positions for the estimated pitches, for the detection of the chord G major.

B. Predetermined Chords—Up to Six Notes

If the number of played notes in a chord is larger than four, the performance of the procedure described in the previous section decays quickly. So, a different procedure based on the definition of a number of chord templates has been developed.

Four different templates have been defined: two templates for major chords and two more for minor chords (see Fig. 13).

For each chord played, all the possible string/fret combinations for all the possible played notes are considered and the template that accounts for the largest number of partials is considered the played chord. In the next lines, a complete example for the chord G major is described in detail.

In Table I, all the peaks in the spectrum of the audio waveform that may correspond to fundamental frequencies of possibly played notes are shown. Now, for each of these notes, all the possible string/fret combinations are evaluated (see Fig. 14).

Each of the templates is evaluated on the whole fingerboard scheme and the total amount of partials is summed for each position. In our example, the chord detected corresponds to the chord G Major with the bass note played on the third fret of the sixth string. Observe, in Fig. 14, that a total amount of 51 partials are accounted for, by the corresponding template at the position indicated (the grayed area in Fig. 14).

VIII. SYSTEM EVALUATION

A number of tests have been done to evaluate the performance of the procedure developed. Two main groups of guitars have been employed in these tests: nine guitar models taken from the RWC database [1] and other four guitar models used for our

TABLE II
AVERAGE INHARMONICITY COEFFICIENTS OF THE SET OF GUITARS USED IN THE EXPERIMENTS

Guitar type	Average inharmonicity coefficients					
	String 1	String 2	String 3	String 4	String 5	String 6
Classical	4.07e-05	6.64e-05	1.29e-04	1.62e-05	1.37e-05	1.83e-05
Acoustic	1.48e-05	4.97e-05	2.77e-05	4.31e-05	6.87e-05	9.92e-05
Electric	1.50e-05	5.02e-05	8.27e-05	5.30e-05	9.04e-05	1.56e-04

own recordings, including classical (nylon string), acoustic and electric types. For each guitar, the inharmonicity coefficients have been estimated for each string. In Table II, the average inharmonicity coefficients obtained for the whole set of guitars tested are shown. These values are obtained analyzing the tuned guitars.

For a complete evaluation of the developments, both single notes and chords have been recorded and analyzed. Note that a comparison with a reference method can not be included since the authors are not aware of other methods published for the joint detection of the note, string and fret for guitar using solely the audio waveform.

The system analyzes audio waveforms sampled at 44.100 kHz. The signal is windowed using a Hann window of 100 ms of duration. When the onset of a note or chord is found, three consecutive windows of length 100 ms are concatenated to apply the procedure described. A fast Fourier transform (FFT) of 2^{18} points is done to the zero-padded signal to calculate the spectrum.

A. Detection Performance on Single Notes

The evaluation of the performance of the system for the detection of the tablature on single notes is based on the analysis of the whole set of 78 notes of the guitar: from the open string note of the 6th string to the 12th fret on the first string. All the notes are played with all the nine guitars of the RWC database [1] and the four models used for our own recordings.

The error rates are defined as the probability of wrong detection of the combination of string/fret/note. Also, string/fret detection error and note detection error are reported separately. Note that the string/fret probability error is defined only for correctly detected notes.

Both the estimated inharmonicity coefficients and the average ones (see Table II) have been used with each guitar, as shown in Table III (note that the recordings labeled EG1 are rather noisy due to the recording conditions; however, the system performs correctly, as it can be observed).

Observe that the error rates are larger for our own recordings mainly because of the quality of the recordings. On the other hand, the error rates are greater when the average coefficients are used instead of the specific coefficients of each guitar, as expected.

The worst results are found for the highest notes of the electric guitar EG132 of the RWC instruments database [1] because, in this case, few partials are detected and in our own recordings. This latter fact is likely due to the presence of noise and the lower quality of the recordings.

TABLE III
ERROR PROBABILITY MEASURED FOR STRING/FRET DETECTION (THE MISLEADING DETECTION OF STRING/FRET IN CASE OF CORRECTLY DETECTED NOTE) AND FOR NOTE DETECTION. BOTH ESTIMATED AND AVERAGE INHARMONICITY COEFFICIENTS ARE USED

Guitar	Error probability				
	Estimated inharm. coef.		Average inharm. coef.		
	string/fret	note	overall	string/fret	note
RWC Instruments Database					
Classical CG091	0.04	0	0.11	0.09	0.03
Classical CG092	0	0	0.01	0.01	0
Classical CG093	0.01	0	0	0	0
Acoustic AG111	0	0	0.06	0.06	0
Acoustic AG112	0	0	0.22	0.22	0
Acoustic AG113	0	0	0.04	0.04	0
Electric EG131	0	0	0.20	0.20	0
Electric EG132	0.01	0	0.26	0.23	0.03
Electric EG133	0	0	0.32	0.32	0
Our own Recordings					
Classical CG1	0.08	0	0.16	0.14	0.03
Classical CG2	0.03	0	0.08	0.05	0.03
Acoustic AG1	0.01	0	0.13	0.13	0
Electric EG1	0.11	0	0.43	0.35	0.09

TABLE IV
RESULTS OF THE EVALUATION OF THE PERFORMANCE OF THE TWO-NOTE CHORDS DETECTION ALGORITHM (ERROR RATE). THE PERCENTAGES OF UNDETECTED NOTES ARE REFERRED TO THE TOTAL NUMBER NOTES PLAYED (128)

Guitar Model	CG1	CG1	AG1	EG1
Mislead string/fret	3.1%	2.3%	1.6%	2.3%
Unplayed note detected	0%	6.2%	1.6%	8.6%
Undetected note	5.5%	3.1%	7.8%	10%

The degradation of the performance of the system is due to the difference between the inharmonicity coefficients used. Note that larger differences can be found among different models of electric guitars than among the other types. Accordingly, as shown in Table III, worst results are found for the electric guitar EG1; the discrepancy between the inharmonicity coefficients of this guitar and the average ones are the greatest.

B. Detection Performance on Free Chords

The performance on the detection of free chords is evaluated using chords of two, three and four notes, played with our own guitars. Note that the number of notes played simultaneously in each chord is also detected by the system.

To begin with, 64 two-note chords are considered. These chords range from simple second to compound fifth. Three types of errors are taken into account, namely: the string/fret is wrongly detected, a played note is not detected and a detected note was not actually played. In Table IV, the results of the evaluation of the performance for the two-note chords are shown.

The performance of the system on the analysis of three-note chords is done using a series of 24 chords, with major and minor examples (with and without inversion). The results obtained are shown in Table V. Note the rise of the error rate of undetected notes with respect to the previous case due to the increase of the complexity of the task.

Finally, four-note chords are considered. 43 four-note chords have been analyzed. As shown in Table VI, the largest error rate

TABLE V
RESULTS OF THE EVALUATION OF THE PERFORMANCE OF THE THREE-NOTE CHORDS DETECTION ALGORITHM (ERROR RATE). THE PERCENTAGES OF UNDETECTED NOTES ARE REFERRED TO THE TOTAL NUMBER OF NOTES PLAYED (72)

Guitar Model	CG1	CG1	AG1	EG1
Mislead string/fret	1.4%	0%	0%	5.6%
Unplayed note detected	1.4%	0%	0%	6.9%
Undetected note	25%	22%	28%	21%

TABLE VI
RESULTS OF THE EVALUATION OF THE PERFORMANCE OF THE FOUR-NOTE CHORDS DETECTION ALGORITHM (ERROR RATE). THE PERCENTAGES OF UNDETECTED NOTES ARE REFERRED TO THE TOTAL NUMBER OF NOTES PLAYED (172)

Guitar Model	CG1	CG1	AG1	EG1
Mislead string/fret	1.7%	1.7%	0%	1.2%
Unplayed note detected	1.7%	0%	0.6%	0.6%
Undetected note	33%	24%	46%	30%

TABLE VII
RESULTS OF THE EVALUATION OF THE PERFORMANCE OF THE PREDETERMINED CHORDS DETECTION ALGORITHM (ERROR RATE). THE PERCENTAGES OF MISLEAD CHORDS ARE REFERRED TO THE TOTAL NUMBER OF CHORDS (36)

Guitar Model	CG1	CG1	AG1	EG1
Mislead chords	8.3%	8.3%	2.8%	0%

corresponds to the undetected played notes, following the trend observed on three-note chords.

As a conclusion, observe that, the performance of the analysis of all the notes detected is good in all cases, with an increasing error rate regarding undetected notes as the number of notes simultaneously played in the chords increases.

C. Detection Performance on Predetermined Chords

Now the performance of the algorithm is evaluated on a series of 36 four-note chords recorded at our laboratory using each of the available guitars. The results are shown in Table VII.

Note that in this case, due to the different behavior of the algorithm, we obtain only a global error rate of chord detection. Also, as it was expected for this detection approach, the error rates obtained are quite lower than the measure that can be found for the detection of four-note chords using the free chord detection algorithm.

Note that some related measures on the detection performance of guitar notes and chords can be found in other works like the ones described in [34] or [35] that confirm the good behavior of our system although all these results cannot be readily compared to the ones obtained by our system since we use different datasets for the evaluation of a rather different target: the identification on note/string/fret combination, instead of the conventional target: the identification of notes and chords played along a guitar piece.

IX. CONCLUSION

A model and schemes for the extraction of the guitar tablature using the audio waveform of guitar notes and chords have been presented in this work.

The model developed is based on the analysis of the inharmonicity coefficient and its influence on the spectrum of the waveform of the notes played with each string on different guitar types.

The inharmonicity coefficients of the guitar to analyze must be estimated beforehand, although predefined sets of coefficients can be used. However, in the latter case, the system will attain a degraded performance.

The detection of the note/string/fret combinations is done without any previous knowledge of the played notes, making use of the analysis of the relation between the fundamental frequency and the partials according the inharmonicity model developed.

The main parameter used in the detection of note/string/fret combinations related to the tablature is the count of the number of partials that follow our model for each possible fundamental frequency detected in the range of the notes that can be played with a guitar.

The procedure developed can also be used iteratively for the detection of chords. In this scenario, the performance of the system decays as the complexity of the chords under analysis grows, as expected. This is the reason why we have developed a template based procedure to analyze chords with more than four notes. A number of tests for the evaluation of the procedures developed for the analysis of notes and chords on the basis of the behavior of the inharmonicity coefficient have been done.

REFERENCES

[1] M. Goto, "Development of the RWC music database," in *Proc. 18th Int. Congr. Acoust. (ICA)*, 2004, vol. I, pp. 553–556.

[2] N. H. Fletcher and T. D. Rossing*, The Physics of Musical Instruments*, 2nd ed. New York: Springer, 1998.

[3] T. D. Rossing*, The Science of String Instruments*. New York: Springer, 2010.

[4] N. T. Rutherford, FINGAR, a genetic algorithm approach to producing playable guitar tablature with fingering instruction Dept. Comput. Sci., Univ. of Sheffield, U.K., 2009, Tech. Rep..

[5] A. Gilardino, "Il problema della diteggiatura nelle musiche per chitarra," *Il "Fronimo"*, vol. 10, pp. 5–12, 1975.

[6] R. Macrae and S. Dixon, "A guitar tablature score follower," in *Proc. Int. Conf. Multimedia Expo*, Jul. 2010, pp. 725–726.

[7] T. D. Rossing, F. R. Moore, and P. A. Wheeler*, The Science of Sound*, 3rd ed. Reading, MA: Addison Wesley, 2002.

[8] I. Barbancho, L. J. Tardon, A. M. Barbancho, and S. Sammartino, "Pitch and played string estimation in classic and acoustic guitars," in *Proc. 126th Audio Eng. Soc. Conv. (AES '09)*, 2009 [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=14897

[9] D. Fragoulis, C. Papaodysseus, M. Exarhos, G. Roussopoulos, T. Panagopoulos, and D. Kamarotos, "Automated classification of piano-guitar notes," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 3, pp. 1040–1050, May 2006.

[10] A. P. Klapuri, "Multiple fundamental frequency estimation based on harmonicity and spectral smoothness," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 6, pp. 804–816, Nov. 2003.

[11] C. Yeh, A. Roebel, and X. Rodet, "Multiple fundamental frequency estimation and polyphony inference of polyphonic music signals," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 6, pp. 1558–7916, Aug. 2010.

[12] T. Gagnon, S. Larouche, and R. Lefebvre, "A neural network approach for preclassification in musical chords recognition," in *Conf. Rec. 37th Asilomar Conf. Signals, Systems, Comput.*, 2003, vol. 2, pp. 2106–2109.

[13] P. D. O'Grady and S. T. Rickard, "Automatic hexaphonic guitar transcription using non-negative constraints," in *Proc. IET Irish Signals Syst. Conf. (ISSC 2009)*, 2009, pp. 1–6.

[14] H. Järveläinen, V. Välimäki, and T. Verma, "Perception and adjustment of pitch in inharmonic string instrument sounds," *J. New Music Res.*, vol. 31, no. 4, pp. 311–319, 2003.

[15] S. Godsill and M. Davy, "Bayesian computational models for inharmonicity in musical instruments," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust.*, 2005, pp. 283–286.
[16] A. S. Galembo and A. Askenfelt, "Study of the spectral inharmonicity of musical sound," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 7, no. 2, pp. 197–203, 1999.
[17] J. Rauhala, H. Lehtonen, and V. Välimäki, "Fast automatic inharmonicity estimation algorithm," *J. Acoust. Soc. Amer.*, vol. 121, no. 5, pp. 1–6, 2007.
[18] A. Radisavljevic and P. Driessen, "Path difference learning for guitar fingering problem," in *Proc. Int. Comput. Music Conf. (ICMC04)*, 2004.
[19] D. Radicioni, C. D. Scienza, and V. Lombardo, "Guitar fingering for music performance," in *Proc. Int. Comput. Music Assoc.*, 2005, pp. 527–530.
[20] S. I. Sayegh and M. F. Tenorio, "Guitar tablature creation with neural networks and distributed genetic search," in *Proc. Conf. Interdisciplinary Musicol.*, 2004.
[21] D. R. Tuohy and W. D. Potter, Advances in applied artificial intelligence ser. Lecture Notes in Computer Science, 2006, vol. 4031/2006, "Generating Guitar Tablature with LHF Notation Via DGA and ANN," in pp. 244–253.
[22] M. Paleari, B. Huet, A. Schutz, and D. Slock, "A multimodal approach to music transcription," in *Proc. 15th IEEE Int. Conf. Image Process. (ICIP)*, 2008, pp. 93–96.
[23] A. M. Burns and M. M. Wanderley, France: IRCAM—Centre Pompidou, "Visual methods for the retrieval of guitarist fingering," in *Proc. NIME '06: Proc. Conf. New Interfaces for Musical Expression*, Paris, France, Jun. 2006, pp. 196–199 [Online]. Available: http://portal.acm.org/citation.cfm?id=1142215.1142263
[24] C. Kerdvibulvech and H. Saito, "Vision-based detection of guitar players' fingertips without markers," *Comput. Graph., Imag., Visualisat. (CGIV)*, pp. 419–428, 2007.
[25] L. L. Wei, Q. A. Salih, and H. S. Hock, "Optical tablature recognition (OTR) system: Using fourier descriptors as a recognition tool," in *Proc. Int. Conf. Audio, Lang. Image Process.*, 2008, pp. 1532–1539.
[26] C. Traube and J. J. O. Smith, "Extracting the fingering and the plucking points on a guitar string from a recording," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust.*, 2001, pp. 7–10.
[27] H. Penttinen, J. Siiskonen, and V. Valimaki, "Acoustic guitar plucking point estimation in real time," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2005, vol. 3, pp. 209–212.
[28] J. Abesser, H. Lukashevich, and G. Schuller, "Feature-based extraction of plucking and expression styles of the electric bass guitar," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2010, pp. 2290–2293.
[29] C. Traube and P. Depalle, "Deriving the plucking point location along a guitar string from a least-square estimation of a comb filter delay," in *Proc. IEEE Can. Conf. Elect. Comput. Eng.*, 2003, pp. 2001–2004.
[30] D. Deutsch, *The Psychology of Music, Second Edition (Academic Press Series in Cognition and Perception)*, 2nd ed. New York: Academic, October 1998.
[31] M. Karjalainen and H. Jarvelainen, "Is inharmonicity perceivable in the acoustic guitar?," in *Proc. Forum Acusticum*, Budapest, Hungary, 2005, pp. 435–440.
[32] H. Fletcher, E. D. Blackham, and R. Stratton, "Quality of piano tones," *J. Acoust. Soc. Amer.* vol. 34, no. 6, pp. 749–761, 1962 [Online]. Available: http://link.aip.org/link/?JAS/34/749/1
[33] S. Lixin, Z. Junxing, and P. Lingbin, "Piano fundamental frequency estimation algorithm based on weighted least square method," in *Proc. Int. Forum Inf. Technol. Applicat., IFITA '09*, May 2009, vol. 2, pp. 155–158.
[34] S. Saito, H. Kameoka, K. Takahashi, T. Nishimoto, and S. Sagayama, "Specmurt analysis of polyphonic music signals," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 3, pp. 693–650, Mar. 2008.
[35] F. Cañadas-Quesada, P. Vera-Candeas, N. Ruiz-Reyes, and J. Carabias-Orti, "Polyphonic transcription based on temporal evolution of spectral similarity of gaussian mixture models," in *Proc. 17th Eur. Signal Process. Conf. (EUSIPCO '09)*, 2009, pp. 10–14.

Isabel Barbancho (SM'10) received the degree in telecommunications engineering and the Ph.D. degree from the University of Malaga (UMA), Malaga, Spain, in 1993 and 1998, respectively, and the degree in piano teaching from the Malaga Conservatoire of Music in 1994.

Since 1994, she has been with the Department of Communications Engineering, UMA, as an Assistant and then Associate Professor. She has been the main researcher in several research projects on polyphonic transcription, optical music recognition, music information retrieval, and intelligent content management. Her research interests include musical acoustics, signal processing, multimedia applications, audio content analysis, and serious games.

Dr. Barbancho received the Severo Ochoa Award in Science and Technology, Ateneo de Málaga-UMA in 2009.

Lorenzo J. Tardon received the degree in telecommunications engineering from University of Valladolid, Valladolid, Spain, in 1995 and the Ph.D. degree from Polytechnic University of Madrid, Madrid, Spain, in 1999.

In 1999 he worked for ISDEFE on air traffic control systems at Madrid-Barajas Airport and for Lucent Microelectronics on systems management. Since November 1999, he has been with the Department of Communications Engineering, University of Malaga, Malaga, Spain. He is currently the head of the Application of Information and Communications Technologies (ATIC) Research Group. He has worked as the main researcher on several projects on audio and music analysis. His research interests include digital signal processing, serious games, audio signal processing, and pattern recognition.

Simone Sammartino received the degree in environmental science and the Ph.D. degree in marine science and engineering from the University of Naples, Naples, Italy, in 2001 and 2006, respectively, and the M.S. degree in telecommunication technology from the University of Malaga (UMA), Malaga, Spain, in 2011.

He joined the Application of Information and Communications Technologies (ATIC) Research Group, UMA, in 2008 Since 2008, he has joined several research projects on music information retrieval and speech analysis. He currently works on music computing, audio signal processing, pattern recognition, and computer graphics.

Ana M. Barbancho received the degree in Solfeo Teaching from the Malaga Conservatoire of Music, Malaga, Spain, in 2000 and the degree in telecommunications engineering and the Ph.D. degree from University of Malaga, Malaga, Spain, in 2000 and 2006, respectively.

Since 2000, she has been with the Department of Communications Engineering, University of Malaga, as an Assistant and then Associate Professor. Her research interests include musical acoustics, digital signal processing, new educational methods, and mobile communications.

Dr. Barbancho was awarded with the Second National University Prize to the Best Scholar 1999/2000 by the Ministry of Education of Spain in 2000 and with the Extraordinary Ph.D. Thesis Prize of ETSI Telecommunications, University of Malaga, in 2007.